Journal of Positive School Psychology http://journalppw.com
2022, Vol. 6, No. 5, 2768-2783

# Prediction Of Learning Disability Of The Children Using Adaptive Effective Feature Engineering Techniques

**C.Radhika[1] and Dr.N.Priya[2]**

[1]*Shrimathi Devkunwar Nanalal Bhatt Vaishnav College for Women, University of Madras, India.*
*radhamegha24@gmil.com*
[2]*Shrimathi Devkunwar Nanalal Bhatt Vaishnav College for Women, University of Madras, India.*

## Abstract

Feature engineering is a critical step in the development of emergent machine learning models. Any process of selection and flexibility is included when using machine learning or mathematical modeling to develop a speculative model. One of the main objectives of predictive modeling is to find a reliable and accurate correlation between a set of available data and a given outcome. Machine learning classifiers are critical for detecting autism spectrum disorders early on. The purpose of this article is to raise awareness about the early detection of ASD in children who are affected. Autism spectrum disorder (ASD) refers to a set of conditions characterised by difficulties with social skills, repetitive activities, speech, and nonverbal communication. We proposed an adaptive CMR-ASD feature engineering model in this study that provides an effective technique for analyzing autism not only for doctors but also for psychologists and learning disability mentors. A hybrid adaptive CMR-ASD model combines various feature selection strategies such as the CHI$^2$test, MUTUAL_INFORMATION, and RFE with PCA to pick a subset of important features by taking into account both the score and ranking of individual features. With appropriate feature selection, this enhanced model is utilized to predict autism in its early stages. A real-time dataset, as well as four different datasets related to autism spectrum disorders, were used in the research. The results showed that the suggested strategy is capable of selecting highly disparate features, and the Matthews correlation coefficient (MCC) is a more reliable statistical rate that generates higher scores than the Cohen's kappa and accuracy scores.

**Keywords :** Autistic, Chi$^2$, MI, RFE, PCA , correlation, MCC

## Introduction

Autism spectrum disorder affects one out of every 44 children in the United States, according to the CDC (ASD). Boys are four times more likely than girls to be diagnosed with autism. An intellectual disability affects 31% of children with ASD, 25% have borderline IQs, and 44% have ordinary to above-average IQs[1]. ASD is a neurobiological disorder influenced by both inherited and environmental factors affecting the developing brain, and there is a list of features related with the risk of having ASD[2]. In a densely populated and demographically 'young' country like India, the frequency of ASD has risen. Differences in assessment methods, however, exist. The authors concentrate on the DSM-5 and the ISAA, two widely used ASD diagnosis instruments in India. The authors discuss the benefits and drawbacks of each of these indicators, as well as suggestions for how to use them in clinical practice[3]. Children with aberrant palmar creases aid in the early detection of neurodevelopment abnormalities such as ASD, allowing for improved outcomes and reducing parental stress and burden [4]. A similar prevalence percentage of 0.09–0.11 was found in both urban and rural Indian populations in a recent comprehensive study of ASD in children aged 0–18 years [5].

## Related Work

In the diagnosis of neuropsychiatric illnesses, machine learning algorithms have begun to show potential. To improve the accuracy of Q-CHAT's categorization, Tartarisco et al. [6], Proposed the use of Artificial Intelligence and machine learning technologies to identify the superlative subset of things capable of accurately differentiating between immature autistic youngsters. Guruvammal, et al. [7], The best features are chosen using the LFCU-LA, which are then fed into a classification technique that uses a hybrid classifier that blends DBN and NN. DBN and NN's hidden neurons' most essential contributions were properly selected, enabling exact identification. Bindu George et al. [8], Robust Kalman filtering-based neural network (RKFNN) is proposed by the author to better correctly detect ASD. RKFNN has been updated to suit the prediction of ASD and has been combined with a neural network for improved outcomes. Schjolberg S, et al. [9], Describes a new M-Chat-R algorithm that can be used to identify autistic children. The M-CHAT-R algorithm's performance was compared to that of the M-CHAT algorithm with proper scoring and ranking. The algorithm's results demonstrate a decline in sensitivity and an increase in specificity when it comes to detecting children with ASD. AK et al. [10], Create new diagnosis software based on video sequences with innovative fuzzy hybrid deep convolutional neural networks with facial expression and gait integration. This technique can be improved by using bio-inspired fuzzy optimizers to reduce dimensionality and increase classification and prediction performance. Goel et al. [11], The MGOA can diagnose Autism Spectrum Disorder at any stage of life. GOA is a nature-inspired algorithm that can successfully explore and exploit the search space. The strategy is numerically tested on three ASD screening datasets aimed at different age groups, including children, adolescents, and adults, and the results are compared to current techniques. Wingfield et al. [12], The author's recommended application is the first to apply an intelligent machine learning model to monitor and detect autism spectrum disease in low- and middle-income nations, using a clinically validated symptom checklist. Machine learning models were constructed and their predictive ability was assessed using clinical pictorial autism assessment schedule data, demonstrating that the random forest was the best classifier for embedding into the mobile screening app.

## Data Collection

The ASD datasets were gathered from the UCI[13], Kaggle[14], and Real-time data repositories, respectively. Based on the categories Toddlers, Child, Adult, Adolescence, and Real-time child data, the datasets are separated into five groups. Real-time data was obtained from the Sree Rehabilitation Center in the Tamilnadu state of India's Chengalpet region. Table[1,2] lists the descriptions of the ASD Dataset and its features. The majority of the attributes in our ASD datasets are 21(besides the infant dataset which has 18 features). Gender, ethnicity, jaundice, own circle of relatives ASD, residence and ASD magnificence are all categorical variables in all datasets, with ten binary traits reflecting the solutions to screening questions (QASD1 to QASD10). Two wide variety factors, inclusive of age and display score/results, also are protected withinside the datasets. The screening questions withinside the infant and child datasets are the same (A1 via A10), but the adolescent and grownup datasets have a few specific questions. In Table [1,2], we have got organized all the surveys from the ASD dataset into 4 categories. Based on the responses to QASD-10 (QASD1 to QASD10) questions, the magnificence score is allotted all through the statistics amassing procedure. When the very last rating of QASD-10 techniques is much less than 6, the magnificence score "No" is assigned. Otherwise, its miles are set to "Yes," indicating that the man or woman certainly has ASD. The cutoff rating for the infant dataset, however, is much less than or the same as 3. As a result, if the whole rating is 4, the affected person is judged to have ASD.

**Table 1: ASD dataset is described in detail.**

| Name | Age | Feature | Instance | Gender | Class | |
|---|---|---|---|---|---|---|
| ASD _Children Dataset | 4-11 yrs | 21 | 292 | M- 207<br>F- 85 | YES<br>NO | 141<br>151 |
| ASD _ Adolescence Dataset | 12-16yrs | 21 | 104 | M-50<br>F- 54 | YES<br>NO | 63<br>41 |
| ASD _ Adult Dataset | >=18 | 21 | 704 | M- 368<br>F- 336 | YES<br>NO | 189<br>515 |
| ASD_Toddlers dataset | 12m – 36m | 18 | 1054 | M-735<br>F-319 | YES<br>NO | 728<br>326 |
| Real-time dataset- | 2-9 yrs | 21 | 82 | M- 48<br>F- 34 | YES<br>NO | 61<br>21 |

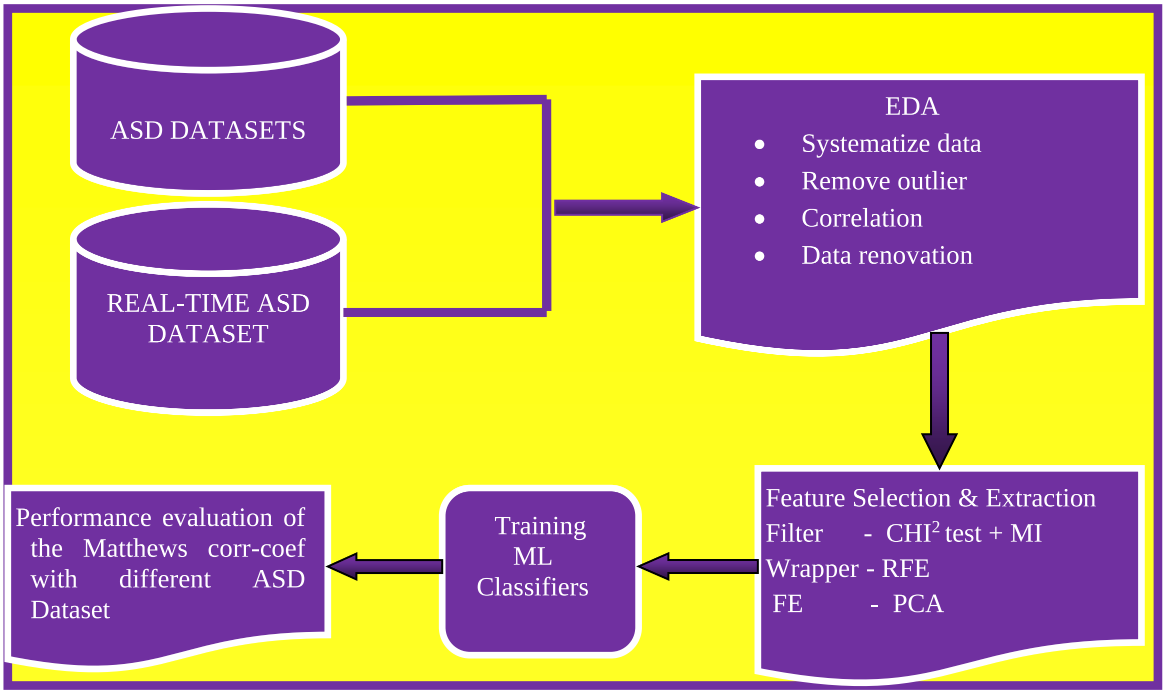
**Table 2: Features of the ASD datasets [15][16].**

| Name of the Attributes | Explication of the Feature |
|---|---|
| QASD1 | • Cognition defects- Inattentive |
| QASD2 | Defective Sociability---Mutual gaze |
| QASD3 | Stereotypical behavior –Intent pattern |
| QASD4 | Cognition defects --Recreation |
| QASD5 | Neurological disorders --pretend |
| QASD6 | Cognition defects —Lack of absorption |
| QASD7 | Cognition defects –No reaction to the query |
| QASD8 | Cognition defects -- Strange tempo |
| QASD9 | Defective Sociability --- Gesticulation |
| QASD10 | Neurological disorders -Atypical visualization. |
| Age | Era |
| Ethnic origins | List of nation |
| Hyperbilirubinemia | Whether or not the case had jaundice when it was born. |
| A member of the family has a history of ASD. | Whether any members of your immediate family have a PDD or have a history of ASD. |
| abode | List of country |
| relationship | Blood relation, Self, therapist, etc |
| Used app earlier | Y/N |
| Score | Integer value based on the attributes QASD1-QASD10 |
| Gender class | F/M |
| Class | N/Y |
| Age-desc | 12m to >=18 era |

## The proposed Adaptive CMR-ASD Methodology

The different ASD datasets have been effectively pre-processed, and now the feature engineering method is being used to identify and extract features. Finally, machine learning models for classification are tested on the reduced set. The Proposed Adaptive CMR-ASD Methodology's Framework is shown in Fig1. It consists of Exploratory Data Analysis(EDA), Reduction of Feature subset, Training ML classifier classifiers.

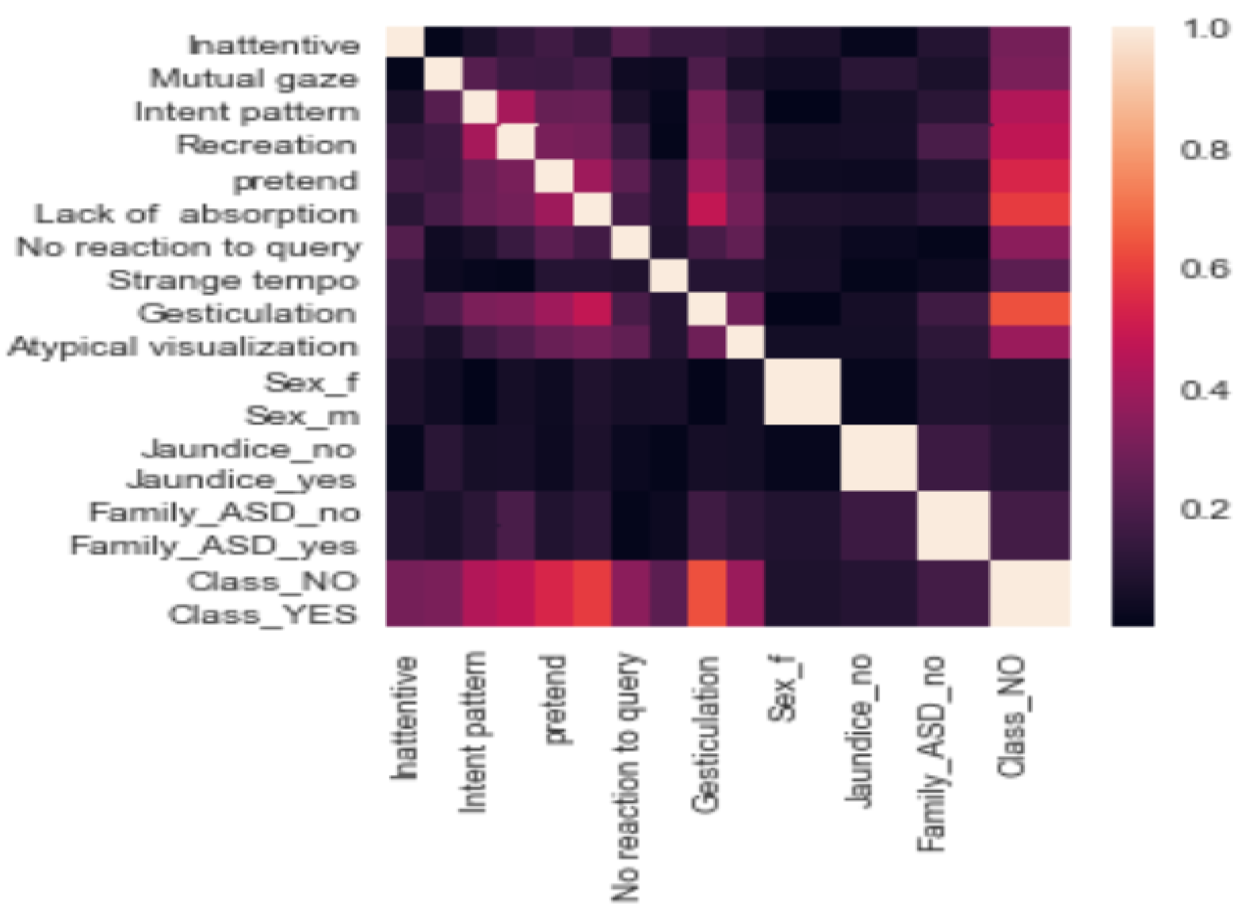
**Fig1: The Proposed Adaptive CMR-ASD Methodology's Framework**




## EDA

The data pre-processing process includes a wide range of activities, including systematizing data, eliminating outliers, correlation, data renovation, and feature extraction-selection of attributes. Utilize the Handle Missing Value technique to transform all missing occurrences of the ASD datasets to the 'unknown' esteem in ASD datasets. As a result, we have excluded some of the variables from the ASD dataset from consideration. For the sake of better analysis, the attributes of ethnicity, who took the exam, and class / ASD features have been removed from the dataset. To convert categorical data into dummy features, after completing the outlier reduction technique, the get dummies method is used after each value of the nominal feature column to turn categorical data into dummy features. To achieve the greatest performance, it is necessary to go through the data renovation procedure, which includes dividing a data set into training and test sets, as well as picking features that are on the same scale. Only one-third of the data will be sent to the training set, with the remaining data being distributed to the test set. The one-hot and label encoder techniques are used to fit and convert the input and target variables. The value of the Pearson correlation coefficient(PCC), on the other hand, reflects the strength of the relationship between the distance variables [17]. Fig2. illustrates PCC between two factors, A PCC more than 90 indicates a significant, positive association between two variables; hence, variables with Pearson correlation coefficients less than 90 are removed from the list.

**Fig 2. Pearson correlation coefficients of ASD dataset**




**Reducing feature subset**

The goal of feature selection is to get rid of any non-essential and/or redundant features, leaving just the ones that are relevant to the scenario. Irrelevant traits can be removed from a person's profile without affecting their learning abilities. The term "redundant features" refers to features that are, in some manner, duplicates of one another. Machine learning is a strategy for picking appropriate attributes for a machine learning techniques based on the type of issue you are seeking to solve [18-20]. There are three types of feature-selection strategies: filter, wrapper, and hybrid. Filtering techniques are the most common. To anticipate the relationship between each independent input variable and each output variable, statistical methodologies are employed in the filter feature selection operations of the filter feature selection procedures. Based on the ranks, chi2, MI, and so forth of intrinsic measures [21-24], this algorithm assigns scores to each attribute. The wrapper model utilizes search strategies such as complete search, random search, and sequential search to choose a subcategory of features from the feature space [25]. To remove excessively redundant items, Hybrid-It applies a filtering strategy. A wrapper strategy is used to offer the additional capabilities that are provided after that. Wrapper time complexity is reduced when a smaller number of attributes are used [25].

**Reducing features of ASD datasets using the Filter method**

Using the chi-square test, we can determine whether or not the categorical data with a given cutoff, such as the threshold values of the expected frequencies e, is more than or equal to the cutoff value of 0.5, or whether or not the categorical data with no cutoff is larger than or equal to the cutoff value of 0. If the e-value is less than the cutoff, then Ho is rejected. If the e-value is less than the threshold, then Ho[26,27] should not be rejected.

Chi-square is symbolized by the symbol

$$chi^2 = \frac{\sum(O_{ij} - Eij)2}{E}$$

------(1)

When filter feature selection methods are used to forecast the target variable, mutual information MI has been effectively used to analyze the relevance and redundancy of a subset of features in predicting the target variable, as well as the redundancy concerning other variables. [28-30]

MI (I, F) is symbolized by the symbol

$$MI(I,F) = entropy(I) - \Sigma S \in vals(F)\frac{Is}{I}\ entropy(Is)$$

------ (2)

The value (F) indicates the attribute F's potential rates, with Is being the subset of I where F has the sum of S. The property I's possible rates are represented by the value (I). Furthermore, the rule of Eq. (2) was the total entropy of I, which was followed by data segregation based on feature F, which was followed by data segregation based on feature F, which was followed by data segregation based on feature F.

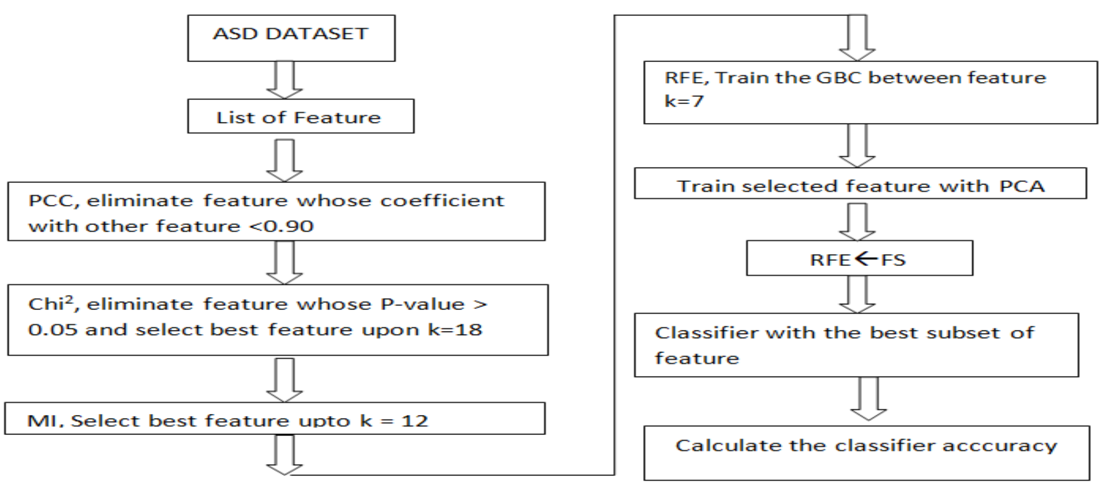
## Reducing features of ASD datasets using the Wrapper method

RFE utilizes several classifiers, each trained on a smaller subset of features (Recursive Feature Elimination). As the number of classifications grows, the time it takes to train a classifier grows as well [31]. Selecting qualities by repeatedly considering smaller and smaller groupings of them is the goal of recursive feature elimination (RFE). First, a small set of features is fed into the estimator, which is then trained on that set. The relevancy of each feature is then computed using any attribute or callable. Then, based on the current collection of characteristics, the least important features are omitted. This procedure is repeated iteratively on the reduced set until the appropriate number of attributes to select from is reached [32]. PCA is used to reduce the dimensionality of a composite dataset, which is then used to extract useful features from the dataset[33].

## Training ML classifier

The proposed Adaptive CMR-ASD Model approach is used in the ASD dataset to discover which features are most relevant for predicting autistic symptoms and which features are not. Results show that the suggested technique is superior to existing algorithms in terms of selecting important attributes. Fig: 3 depicts the structural design of the proposed Adaptive CMR-ASD Model feature selection. Table: 5 indicates the most significant reductions in the final features set for different ASD datasets.

**Fig. 3. FRAMEWORK OF PROPOSED ADAPTIVE MODEL**



## Feature Engineering

In this proposed adaptive model, the effective data pre-processing process encompasses a wide variety of activities such as data systematization, reducing outlier, correlation, and data renovation. Following this feature selection, filter and wrapper-based approaches are demonstrated. In phases I and II of the pseudo-code description of CMR-ASD model are exposed in Table [3,4], the selectkbest technique is used to extract the best features based on the k highest score using various statistical approaches. To choose the best feature with a high ranking score, we employ filter-based feature selection methods such as $chi^2$ and MI. We may now reduce the number of features in the ASD dataset by a factor of two by using $chi^2$ and the ideal subset based on the highest-ranking score, i.e. k=18. Following that, MI undergoes a decreasing feature subset procedure to produce the ideal feature subset based on the high ranking score, in this case, k=12. Following that, in the wrapper approach, the reduced subset is used to RFE to provide superior adaptive feature selection methods. In this situation, k=7, RFE is subjected to a decreasing feature subset procedure to produce the ideal feature subset based on the high ranking score. Table: 5 lists the significant characteristics that were considered for inclusion in the classification procedure for the ASD dataset on the interaction, imagination, gestures, conversation, Jaundice, and vital aspects are picked mostly based on verbal exchange, communiqué, and facial features. Following the estimation of feature significance, PCA is used to train the reducing feature subset and construct a machine learning Classifier using the proposed Adaptive CMR-ASD feature engineering approach. The performance assessment and metrics of several machine learning classifiers are analyzed in detail. The graphical depiction of the feature importance of each ASD dataset is shown in Fig:[3-4].

**Table: 3 Phase I— CMR-ASD model**

```
Step1: Begin
Step2: Df1← ASD_dtset (df1)     //EDA
Step3: Calculate the Pcc between the features in df1
Step4: If  Pcc < 0.90
            eliminate feature
        Else
            FS← Pcc
        End if
Step5: tr_In (Xtr, Xts)
        apply one hotencoder ()
        Fit and Transform the data
        return  Xtr, Xts
Step6: tr_tar (Ytr,Yts)
        apply labelencoder ()
        Fit and Transform the data
        return (Ytr, Yts)
Step7: set_FS (Xtr, Ytr)
        FS[]// Filter method
Step8: For  Fi ← 1 to F each
            Calculate the Chi2 between features  in df1
Step8a:   If p >= 0.5 and k <= 18
              append   Fs← Fi
        End for
```

```
    // Rank FS according to MI with df1
Step9: For  Fj ← 1 to FS each
           MI(Fj) ← (MI(Xtr,Ytr), K<=12)
           append FS← MI(Fj)
          End for
Step10: End
```

**Table: 4 Phase II— CMR-ASD model**

```
Input    df1→ Original ASD data set
         F→ Set of Features
Output  FS→ Optimal Feature Subset.

Step1: Begin
Step2: Load ASD_dtset(df1)
Step3: Xtr, Ytr ← ASD_dtset(ASD)
Step4: While (df1 <> null ) do
Step5: splitting the data set into training & testing
Step6: tr_I ← tr_In(Xtr, Xts)
Step7: tr_t ←tr_tar(Ytr, Yts)
Step8: FS, XtrFS, XtrFS  ← sel_FS(tr_I, tr_t) //wrapper method
Step9: rfe ←Train the GBC between the feature in df1
Step10: Compute the weight vector of the GBC
            Rfe ← Grfe(XtrFS, XtrFS )
            Find the bottom _rank feature
            Grfe ←Grfe-bottom_rank feature
            GFS ← Grfe
Step11: Train GFS  with PCA
Step12: Train ML classifiers to reduce GFS
Step13: Construct the wgt vector of GRFFS(GAFS1,..., GAFSw)
Step14: Calculate the classification accuracy with MCC
Step15: End while
Step16: End
```

**Table:5 Reduce feature subset of ASD datasets.**

| ASD Dataset | Reduced feature subset |
|---|---|
| TODDLERS | QASD1, QASD2, QASD7, QASD8, QASD9, QASD10, Jaundice |
| CHILD | QASD1, QASD6, QASD7, QASD8, QASD9, QASD2, Jaundice |
| ADOLESCENT | QASD1, QASD3, QASD4, QASD6, QASD9, QASD10, Jaundice |

| | |
|---|---|
| ADULT | QASD6, QASD5, QASD7, QASD4, QASD9, QASD10, Jaundice |
| REAL-TIME | QASD1, QASD2, QASD9,Jaundice, QASD6, QASD7, QASD8 |

**Fig:3 The graphical depiction of the feature relevance of each Child, Adult dataset.**

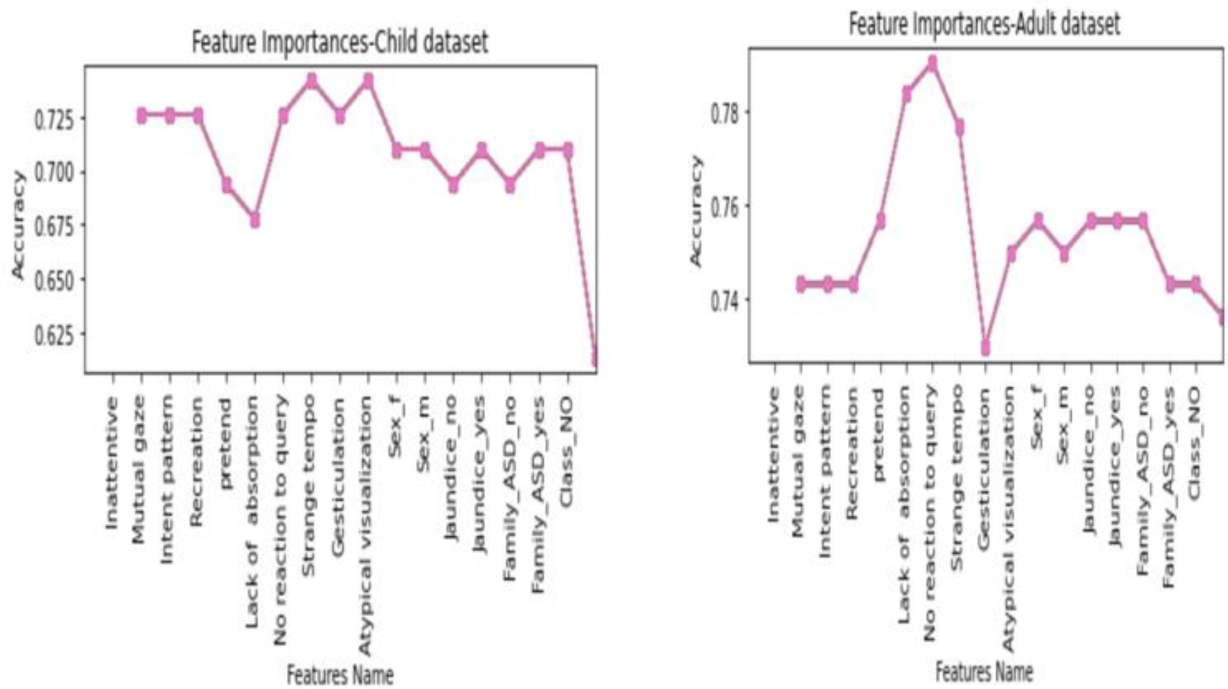
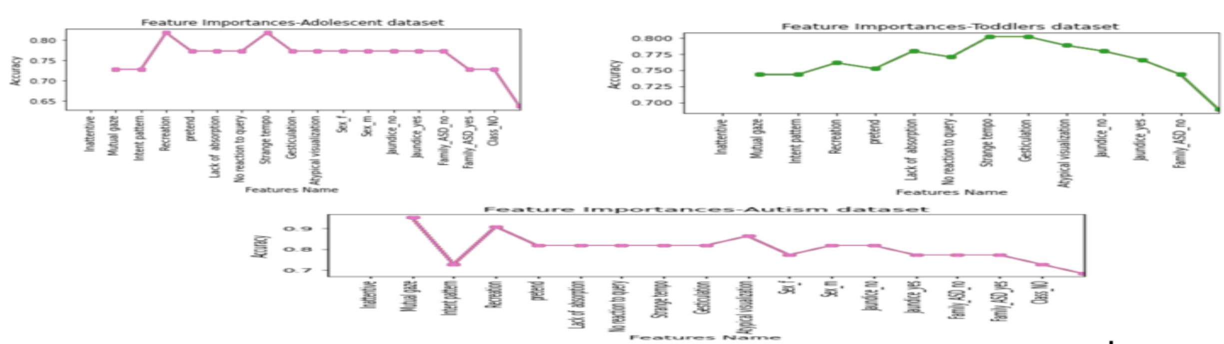



**Fig:4 The graphical representation of the feature importance of each Adolescent, Toddler, Real-time Autism dataset.**




## Experimental Result and discussion

A full systematic analysis and performance validation of the proposed adaptive CMR-ASD system are provided in this segment. The proposed adaptive CMR-ASD model was evaluated against a variety of ASD datasets to determine its overall performance. Effective data preprocessing approaches and dimensionality reduction strategies for feature selection (Filter + wrapper based methods) and extraction – PCA with Random Forest are used in this work to predict autistic data. Since the majority of the characteristics in our ASD datasets are 21 (except for the newborn dataset, which has 18 features), the proposed adaptive CMR-ASD model is applied to pick the seven best features from the majority of the attributes in the ASD datasets. Among the seven characteristics is jaundice, which indicates that children born with a high incidence of jaundice are more likely to develop an autistic disorder than other children. Fig 3-6 illustrates the effect of the ASD dataset's most salient characteristic on the Real-time dataset. The proposed adaptive CMR-ASD feature engineering approach was validated in terms of assessment metrics, prognostic accuracies, and diagnostic plot performance analysis in comparison with SVM, cart machine learning algorithms, and it was discovered to be propounding satisfying results as shown in Table:6.

**Table:6 The proposed adaptive model improves the performance accuracy of several ASD datasets.**

| ASD Dataset | Proposed adaptive model | SVM | CART |
|---|---|---|---|
| ASD_Toddlers dataset | 0.98 | 0.77 | 0.93 |
| ASD _Children Dataset | 0.99 | 0.76 | 0.92 |
| ASD _ Adolescence Dataset | 1.0 | 0.79 | 0.97 |
| ASD _ Adult Dataset | 0.97 | 0.81 | 0.90 |
| Real-time Dataset | 0.95 | 0.85 | 0.90 |

### Metrics for classifier assessment

Multiple evaluation measurements for the proposed adaptive model are presented in Table [8-12] based on the confusion matrix, k-fold cross-validation, and different correlation coefficients. The confusion matrix, k-fold cross-validation, and different correlation coefficients are all represented mathematically in Table:7. If we examine a binary classification issue for autism,

True positives(tp) are those that accurately predict an autistic patient's behavior, False negatives (fn) are actual positives that are incorrectly projected non-autistic patients, True negatives(tn) are negative results that are seen in patients who are not autistic, and False positives(fp) are actual negatives that are incorrectly anticipated, autistic patients. The following assessment measures were utilized to evaluate the findings: precision, recall, f1, acc, ERR MCC, and Cohen's kappa statistic. [34] The recall metric is the proportion of relevant data to the total data retrieved. Precision is defined as the ratio of correct data to retrieved data. The system's F-value is defined as the weighted harmonic mean of its accuracy and recall. Accuracy is defined as the percentage of data instances properly categorized over the total data instances. Because accuracy and error rate are inversely proportional, we can always get the other by subtracting the other.

**Table:7 Metrics for classifying data are defined [35-41]**

| | |
|---|---|
| F1 | $2 * \frac{P * R}{P + R}$ |
| Precision(P) | $\frac{tp}{tp + fp}$ |
| Recall(R) | $\frac{tp}{fn + tp}$ |
| ERR | $\frac{fp + fn}{fp + fn + tp + tn}$ |
| Acc | 1-ERR |
| MCC | $\frac{tp * tn - fp * fn}{\sqrt{(tp + fp)(tp + fn)(tn + fp)(tn + fn)}}$ Worst case=-1; Best case =1 |

| Cohen's Kappa | $\frac{2 * (tp.tn - fp.fn)}{(tp + fp)(fp + tn) + (tp + fn)(fn + tn)}$ |
|---|---|

MCC appears to be the most promising option since it is not only interpretable but is also resistant to changes in the prediction target [35-37]. It is the binary classification rate that provides a high score only when the binary predictor successfully predicted the majority of positive and negative data instances in the training data set. The Kappa statistic is a conformance metric for categorical data that compares agreement to what would be predicted by chance. Cohen's Kappa [38-41] focused on assessing agreement between two observers who rated the same group of individuals on a nominal scale with two or more classes. Additionally, the measure is frequently employed in two-class classification. When compared to Cohen's kappa, accuracy, and F1 score, among other measures, the MCC produces a more accurate and informative result.

**Table: 8 Accuracy of performance on ASD_Toddlers dataset using the proposed adaptive CMR-ASD model**

| Evaluation metrics | F1_score | Recall | Precision | Error-rate | Cohen_kappa | Matthews_corrcoef |
|---|---|---|---|---|---|---|
| Proposed method | 0.98 | 0.95 | 0.94 | 0.02 | 0.87 | 0.92 |
| SVM | 0.79 | 0.78 | 0.80 | 0.23 | 0.71 | 0.78 |
| CART | 0.94 | 0.92 | 0.97 | 0.07 | 0.85 | 0.90 |

**Table: 9 Accuracy of performance on ASD_ Children dataset using the proposed adaptive CMR-ASD model**

| Evaluation metrics | F1_score | Recall | Precision | Error-rate | Cohen_kappa | Matthews_corrcoef |
|---|---|---|---|---|---|---|
| Proposed method | 0.98 | 1.00 | 0.96 | 0.01 | 0.85 | 0.91 |
| SVM | 0.86 | 1.00 | 0.75 | 0.24 | 0.75 | 0.80 |
| CART | 0.95 | 0.94 | 0.96 | 0.08 | 0.90 | 0.95 |

**Table: 10 Accuracy of performance on ASD_ Adolescent dataset using the proposed adaptive CMR-ASD model**

| Evaluation metrics | F1_score | Recall | Precision | Error-rate | Cohen_kappa | Matthews_corrcoef |
|---|---|---|---|---|---|---|
| Proposed method | 1.00 | 1.00 | 1.00 | 0.0 | 0.84 | 0.92 |
| SVM | 0.86 | 1.00 | 0.76 | 0.21 | 0.71 | 0.75 |
| CART | 0.98 | 1.00 | 0.96 | 0.03 | 0.83 | 0.88 |

**Table:11 Accuracy of performance on ASD_ Adult dataset using the proposed adaptive CMR-ASD model**

| Evaluation metrics | F1_score | Recall | Precision | Error-rate | Cohen_kappa | Matthews_corrcoef |
|---|---|---|---|---|---|---|
| Proposed method | 0.97 | 0.94 | 0.95 | 0.03 | 0.88 | 0.91 |
| SVM | 0.85 | 0.98 | 0.75 | 0.19 | 0.73 | 0.75 |
| CART | 0.95 | 0.92 | 0.98 | 0.10 | 0.86 | 0.89 |

**Table:12 Accuracy of performance on Real-time dataset using the proposed adaptive CMR-ASD model**

| Evaluation metrics | F1_score | Recall | Precision | Error-rate | Cohen_kappa | Matthews_corrcoef |
|---|---|---|---|---|---|---|
| Proposed method | 0.98 | 0.94 | 0.99 | 0.05 | 0.90 | 0.95 |
| SVM | 0.85 | 0.98 | 0.75 | 0.15 | 0.60 | 0.65 |
| CART | 0.96 | 0.92 | 0.92 | 0.10 | 0.88 | 0.90 |

We evaluate the performance of our proposed technique with that of the current benchmark methods, Umamaheswari et al. [42] create an IWOA-FRBC model that is used to determine the different class labels of ASD. Three benchmark ASD datasets are used to examine the effectiveness of the IWOA-FRBC model. By achieving the greatest accuracy of each dataset, the obtained simulation results demonstrated the integrity of the IWOA-FRBC model. As shown in Table [compare], the accuracy of those datasets is 93%, 95%, and 94% for the child, adolescent, and adult datasets, respectively, according to the authors. The model presented by Kazi Shahrukh Omar et al.[43] for the prediction of autism, features included random forest-ID3 and random forest-CART. The AQ10 dataset and 250 actual datasets were used in the evaluation. The accuracy, precision, and false-positive rate of the tests were compared. There is 92% accuracy for children, 93% for adolescents, and 97% for adults in those datasets, as shown in Table: 13. With an accuracy rate of 97%, the logistic regression framework of [44] Vakadkar et.al was able to identify essential aspects of the dataset for autism screening in toddlers. In the proposed model, we eliminated highly correlated variables, processed them, and applied various machine learning classifiers, which produced better results than others. Table: 13 presents a comparison of current works with the proposed model, demonstrating that the proposed model achieves much better results than earlier studies on a variety of assessment measures.

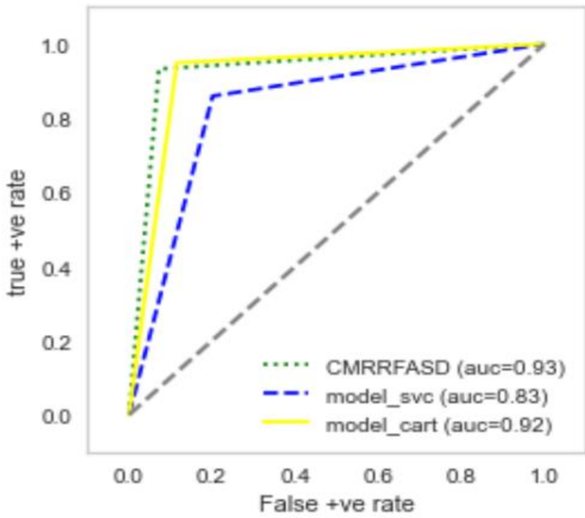
**Table:13 Compare the accuracy of the forecast with the work of other authors.**

| Ref | Model | Accuracy | | | | |
|---|---|---|---|---|---|---|
| | | Child | Adolescent | Adult | Toddlers | Real-time |
| [35] | IWOA-FRBC model | 93% | 95% | 94% | - | - |
| [37] | LR | - | - | - | 97% | - |
| [36] | Merging RF-CART and RF-ID3 | 92% | 93% | 97% | - | |
| | Proposed Adaptive CMR-ASD | 99% | 100% | 97% | 98% | 95% |

In the final analysis, we depict the receiver operating characteristic (ROC) curve for the best classifier in ASD datasets, which turned out to be a random forest. As we can see from the ROC curves in Fig[5-9], we can recognize ASD with a moderate to high true positive rate, which is more than that of the preliminary test, while maintaining an equal false positivity rate.

**Fig:5 The ASD_Toddlers dataset ROC**

**Fig: 6 The ASD_Childern dataset ROC**



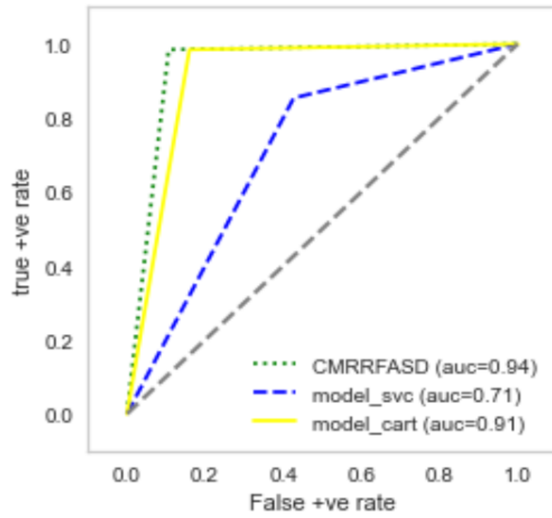
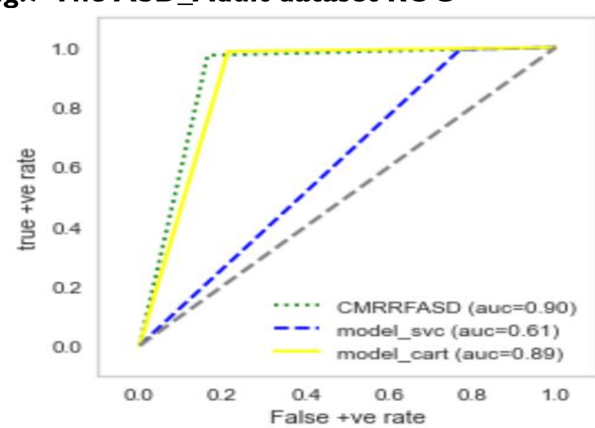




**Fig:7 The ASD_Adult dataset ROC**

**Fig:8 The ASD_Adolescent dataset ROC**



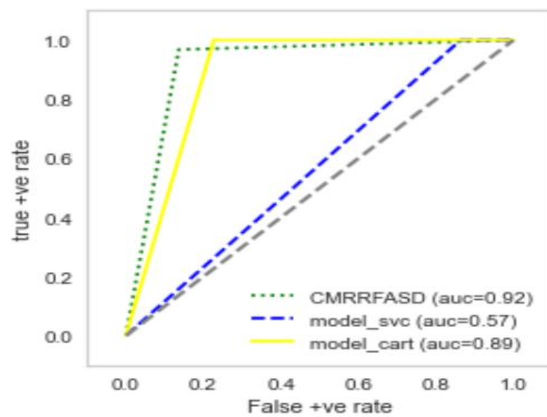
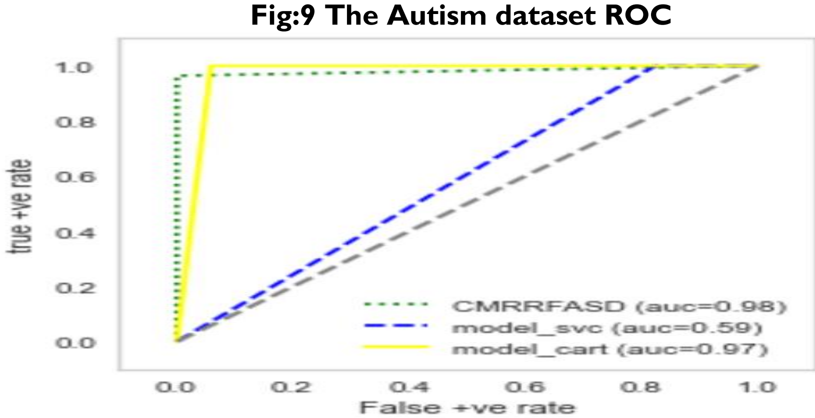




**Fig:9 The Autism dataset ROC**




## Conclusion

For clinical conclusion, machine learning provides a huge step forward in the usage of easily available information as a tool for development and progress. Adaptive CMR-ASD is more efficient than previous machine learning algorithms. The CMR-ASD approach is used to categorise the five distinct autism datasets. In addition to EDA, the proposed adaptive model uses dimensional reduction techniques including filter and wrapper selection, as well as scoring and rating feature subset extraction on RF. The Matthews correlation coefficient (MCC) is more reliable than accuracy, F1, and Cohen's kappa. The obtained simulation result indicated that the CMR-ASD model produced efficient results, with a 98% accuracy on Toddlers, 96% on children, 1.0% on adolescents, 97% on Adults, and 95% on Real-time datasets respectively. In this study, we examined the effectiveness of the proposed adaptive CMR-ASD model in the prediction of autism spectrum disorder specifying that kids born with Cognition defects, Defective sociability, Jaundiced have autism disorder more possibility than other kids. Due to cultural considerations, knowledge of ASD is low in economically underdeveloped countries. Patients with ASD are typically kept untreated for lengthy periods due to a lack of resources. Our next study will focus on gathering more data from other sources and enhancing the suggested AI classifier's accuracy. The results show the classifier's behaviour as the feature is reduced. These findings can also help ASD persons gain access to essential emotional support networks that can help them succeed in the future.

## Reference

1. Maenner MJ, Shaw KA, Bakian AV, et al. Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years- Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2018. MMWR Surveill Summ 2021;70(No. SS-11):1-1. DOI: http://dx.doi.org/10.15585/mmwr.ss7011a1external icon.
2. Hodges, H., Fealko, C., & Soares, N., Autism spectrum disorder: definition, epidemiology, causes, and clinical evaluation. Translational Pediatrics, 2020. 9(Suppl 1), S55-S65. https://doi.org/10.21037/tp.2019.09.09
3. Samir H. D, Deepti K. M, Ameya P.B, Diksha G. Analysis of Tools for Diagnosing Autism Spectrum Disorder in the Indian Context. 004 Acard J Ped Neonatol. 2016; (3):555562. DOI: 10.19080/AJPN.2016.01.555562.
4. Avni Gupta, Aakanksha kharb, Dr. Sujata Sethi. Study of Abnormal Palmer Creases in Children with Autism spectrum disorder: A case-control study in north india. Paripex-Indian Journal of Research: vol-10 | issue-5| May-2021.
5. Chauhan A, Sahu JK, Jaiswal N, Kumar K, Agarwal A, Kaur J, et al. Prevalence of autism spectrum disorder in Indian children: A systematic review and meta-analysis. Neurol India 2019;67:100-4
6. Tartarisco, G., Cicceri, G., Di Pietro, D., Leonardi, E., Aiello, S., Marino, F., Chiarotti, F., Gagliano, A., Arduino, G.M., Apicella, F., Muratori, F., Bruneo, D., Allison, C., Cohen, S.B., Vagni, D., Pioggia, G., Rutta, L.: Use of Machine Learning to investigate the Quantitative Checklist for Autism in Toddlers(Q-CHAT) towards Early Autism Screening. Diagnostics. Vol. 11(3), pp.574 (2021), http://doi.org/10.3390/diagnostics11030574
7. Guruvammal S, Chellatamilan T and Jegatha Deborah L. Optimal Feature Selection and Hybrid Classification for Autism Detection in Young Children, Computational Intelligence, Machine Learning and Data Analytics the Computer Journal, Vol.00 no.00,2020.
8. Bindu George, Dr Chandra Blessie E. Autism Spectrum Disorder Prediction Using Robust Kalman Filtering Based Neural Network. Journal of Theoretical and Applied Information Technology 2021. Vol 99. No 11.
9. Schjolberg S, Shic F, Volkmar FR, Nordahl-Hansen A, Stenberg N, Torske T, Larsen K, Riley K, Sukhodolsky DG, Leckman JF, Chawarska K, Oien RA. What are we optimizing for in autism screening? Examination of algorithmic changes in the M-CHAT. Autism Res. 2022 Feb;15(2):296-304. doi:10.1002/aur.2643. Epub 2021 Nov 26. PMID:34837355; PMCID:PMC8821132.
10. AK, Saranya, Anandan, R..FIGS-DEAF: an novel implementation of hybrid deep learning algorithm to predict autism spectrum disorders using facial fused gait features. Distributed and Parallel Databases 2021. 10.1007/s10619-021-07361-y.
11. Goel, Nikita; Grover, Bhavya; Anuj,; Gupta, Deepak; Khanna, Ashish; Sharma, Moolchand. Modified Grasshopper

Optimization Algorithm for detection of Autism Spectrum Disorder. Physical Communication 2020. 41(), 101115-doi:10.1016/j.phycom.2020.101115.

12. Wingfield, Benjamin; Miller, Shane; Yogarajah, Pratheepan; Kerr, Dermot; Gardiner, Bryan; Senevirathe, Sudarshi; Samarasinghe, Pradeepa; Coleman, Sonya. A predictive model for paediatric autism screening. Health Informatics Journal 2020, doi:10.1177/1460458219887823.
13. UCI Machine Learning Repository. Retrieved http://Archive.Ics.Uci.Edu/ML/Index.Php
14. Kaggle Repository. Retrieved https://www.kaggle.com/fabdelja/autism-screening-for-toddlers
15. Thabtah, F.: An accessible and efficient autism screening method for behavioral data and predictive analyses. Health informatics J. Vol 25(4), pp.1739-1755, 2019. Doi:10.1177/1460458218796636.
16. Thabath, F.: ASDTests. A mobile app for ASD screening. [Internet]. 2017[cited 2018 Dec 20]. Available from www.asdtests.com.
17. Pengtian Chen et al. Research on Intrusion Detection Method Based on Pearson Correlation Coefficient Feature Selection Algorithm , 2021. J. Phys.: Conf. Ser. 1757 012054.
18. Said Bahassine, Abdellah Madani, Mohammed Al-Sarem, Mohammed Kissi. Feature selection using an improved Chi-square for Arabic text classification. Journal of King Saud University- Computer and Information Sciences, Vol-32(2), 2020. Pp: 225-331, ISSN 1319-1578, https://doi.org/10.1016/j.jksuci.2018.05.010.
19. Alhaj TA, Siraj MM, Zainal A, Elshoush HT, Elhaj F. Feature Selection using Information Gain for improved structural-Based Alert Correlation. PLoS ONE 2016. 11(11):e0166017. Doi:10.1371/journal.pone.0166017.
20. Washington, P., Paskov, K.M., Kalantarian, H., Stockham, N., Voss, C., Kline, A., Patnaik, R., Chrisman, B., Varma, M., Tariq, Q., Dunlap, K., Schwartz, J., Haber, N., & Wall, D. P. Feature Selection and Dimensional Reduction of Social Autism Data. Pacific Symposium on Biocomputing 2020, 25,707-718.
21. Shroff, K,P., Maheta, H,H.: A comparative study of various feature selection techniques in high-dimensional dataset to improve classification accuracy. International conference on computer communication and informatics. pp. 1-6 (2015). doi: 10.1109/ICCCI.2015.7218098.
22. Huan Liu., Lei yu.: Toward integrating feature selection algorithms for classification and clustering. IEEE Transactions on knowledge and data engineering. vol. 17, pp. 491-502 (2005). doi: 10.1109/TKDE.2005.66.
23. Dash, Manoranjan., Huan Liu.: Consistency-based search in feature selection. Artificial intelligence 151. pp.155-176 (2003).
24. Dash, Manoranjan., Huan Liu,: Feature selection for classification. Intelligent data analysis1. pp. 131-156 (1997).
25. Kohavi, R., George, H.J.: Wrappers for feature subset selection. Artificial intelligence 97. pp. 273-324 (1997).
26. Chi-Square test for Independence. (2021, January 11). CoConino Community College. https://stats.libretexts.org/@go/page/5227
27. Bachri, O.S. Feature Selection based on Chi Square in Artificial Neural Network to predict the accuracy of student. International Journal of Civil Engineering and Technology. Vol -8(8) 2017, pp. 731-739, ISSN Print:0976-6308 and ISSN Online: 0976-6316
28. Wang, Yan & Cang, Shuang & Yu, Hongnian. Mutual Information Inspired Feature Selection Using Kernel Canonical Correlation Analysis. Expert Systems with Application: X.4. 100014. doi.org/10.1016/j.eswax.2019.100014
29. Prasetiyowati, M.I., Maulidevi, N.U. & Surendro, K. Determining threshold value on information gain feature selection to increase speed and prediction accuracy of random forest. J Big Data V01-8(84) 2021. doi.org/10.1186/s40537-021-00472-4
30. Lei, S. A Feature Selection Method Based on information Gain and Genetic Algorithm. 2012 International Conference on Computer

Science and Electronics Engineering, 2, 355-358.

31. Guyon, I., Weston, J., Barnhill, S., Vapnik, V.: Gene selection for cancer classification using support vector machine, Machine Learning, vol. 46, pp. 389-422 (2002).
32. Pedregosa et al., Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research, vol.12, pp. 2825-2830(2011).
33. Sebastian Raschka, Python Machine Learning, (2015), ISBN: 978-1-78355-513-0, www.packtpub.com
34. Sofia Visa, Brian Ramsay, Anca Ralescu, Esther van der Knaap, Confusion Matrix-based Feature Selection, CEUR Workshop Proc., vol 710, pp. 120-127, 2011.
35. Chicco, D., Jurman, G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation, BMC Genomics 21, 6(2020). https://doi.org/10.1186/s12864-019-6413-7.
36. Delgado R, Tibau X-A, Why Cohen's kappa should be avoided as performance measure in classification. PloS ONE 2019, 14(9): e0222916. https://doi.org/10.1371/journal.pone.0222916.
37. D. Chicco, M. J. Warrens and G. Jurman, The Matthews Correlation Coefficient (MCC) is more informative than Cohen's kappa and Brier score in binary classification assessment, IEEE Access, vol. 9, pp. 78368-78381,2021, doi: 10.1109/ACCESS.2021.3084050.
38. M. J. Warrens, Cohen's kappa can always be increased and decreased by combining categories, Stat.Methodol., vol. 7, no. 6, pp. 673-677, nov. 2010.
39. J. Cohen, A coefficient of agreement for nominal scales, Educ. Psychol. Meas., vol. 20, no. 1, pp. 37-46, Apr. 1960.
40. M. J. Warrens, On the equivalence of Cohen's kappa and the Hubert-Arabie adjusted Rand index, J. Classification, vol. 25, no.2, pp. 177-183, nov. 2008
41. H. C. Kraemer, Kappa coefficient, in Wiley Stats Ref: statistics reference online. New York, NY, USA: Wiley,2014, pp.1-4.
42. Umamaheswari K, Dr. Latha Parthiban. Improved Whale Optimization Algorithm Based Feature Selection with Fuzzy Rule Classifier for Autism spectrum disorder diagnosis. International Journal of Advanced Research in Engineering and Technology 2020. Vol-11, Issue-11, pp.41-55, ISSN Print: 0976-6480 and ISSN Online:0976-6499. doi: 10.34218/IJARET.11.11.2020.005.
43. K. S. Omar, P. Mondal, N. S. Khan, M. R. K. Rizvi and M. N. Islam. A Machine Learning Approach to Predict Autism Spectrum Disorder. 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), 2019, pp. 1-6, doi: 10.1109/ECACE.2019.8679454.
44. Vakadkar, K., Purkayastha, D., & Krishnan, D. (2021). Detection of Autism Spectrum Disorder in Children Using Machine Learning Techniques. SN Computer Science, 2(5). doi:10.1007/s42979-021-00776-5.